

GitHub 習熟度評価手法の提案とシステム開発 PBL における習熟度評価の試み

伊藤 恵 松原 克弥 富永 敦子

システム開発 PBL において、多くの学生チームは GitHub などの標準的なツールを利用している。我々はこのツールの利用履歴から学生たちの PBL 活動の客観的かつ定量的な評価を行うことを目指している。GitHub の利用履歴データを用いた学生チームの活動評価を目指していたが、個々の学生の GitHub に対する習熟度の差異が大きく、チーム全体の評価が困難となった。そこで我々は、個々の学生の GitHub 習熟度を評価するために、チェックリストを作成し、実際にシステム開発 PBL に参加している学生を対象として評価実験を行った。本稿ではチェックリストに基づく GitHub 習熟度評価手法とその評価実験の結果について述べる。

In system development PBLs, many student teams use several standard tools, such as GitHub. By usage history data of such tools, we aim to evaluate objectively and quantitatively students' activities in PBL. However, evaluation using usage history data of such tools is greatly affected by proficiency level of students. So, it is also needed to evaluate of proficiency level of students. For that reason, we create a checklist for proficiency level of GitHub and try to evaluate proficiency level by the checklist in an on-campus Hackathon event and PBL. In this article, we report the checklist and the evaluation results by the checklist.

1 はじめに

さまざまな教育分野において実践的な教育方法の一つとして Project Based Learning(以下 PBL)が行われており、情報系の学部や学科においても、実践的な IT 教育の一つとしてシステム開発 PBL が広く実施されるようになってきている。[1][2] このようなシステム開発 PBL における成績評価は発表会や最終成果物に対する評価が中心である。しかし、学習対象には発表や最終成果物だけでなく、そこに至るまでのプロセスが含まれていることも多い。しかし、システム開発 PBL において開発プロセスの評価方法は確立されているとは言えない。[3]

システム開発 PBL では GitHub などの定番ツールがよく使われる。GitHub^{†1} はシステム開発等で使

われるバージョン管理ツールの一つ Git を基にした Web 上のサービスであり、開発プロセスに紐づけられる利用履歴情報が相応に残される。我々はその情報に着目し、GitHub を使用するシステム開発 PBL を対象に、利用履歴情報を用いた PBL プロセス評価手法の確立を目指している。[5] しかし、Git/GitHub はまあまあ使えるレベル、かなり使いこなしているレベルなど、使用者の習熟度合の幅が広く、システム開発 PBL においてもチーム間やチーム内で習熟度合の差が大きくなることも多く、それが利用履歴を用いたプロセス評価に大きく影響してしまう。

そこで我々は、システム開発 PBL のプロセス評価に活用することを前提として GitHub の習熟度チェックリストを作成し、実際にこれを用いた習熟度の評価実験を行うこととした。

2 主観評価に基づく習熟度評価

最終的には客観的かつ定量的な評価を目指しているが、Git や GitHub の習熟度の基準やそれを評価するための評価項目が確立されていないことから、まず

Proposing a Method to Evaluate Learning Level of GitHub and a Trial Evaluation.

Kei Ito, Katsuya Matsubara, Atsuko Tominaga, 公立はこだて未来大学, Future University Hakodate.

^{†1} <https://github.com>

はアンケート調査に基づく評価方法の確立を目指す。ここでは、GitHub 習熟度評価のためのアンケート項目とアンケート結果からの習熟度判定方法について述べる。

2.1 アンケート項目

アンケート項目は、Git や GitHub に関する書籍やある程度整理された Web 上の情報[7][4]を参考に項目候補を決め、複数の GitHub 熟練者で協議した上で決定した。表 1 がアンケート項目の一覧である。表に記載の項目順に、それぞれの項目について「何のことか分からない」「コマンドは分かる/使ったことはある」「分かる/使っている」「使いこなしている/十分理解している」の 4 択で回答させる。表中の「分類」は、Git の基本的な操作に関するもので、Git/GitHub の利用開始初期から使うと思われるものを「基本 1」、基本的な操作だが利用開始後しばらくしてから使うと思われるものを「基本 2」とした。また、GitHub Flow などのバージョン管理のワークフローにおいて多く使われるブランチ操作についても同様に、ブランチの利用開始初期から使うと思われるものを「ブランチ 1」、しばらくしてから使うと思われるものを「ブランチ 2」とした。

ただし、これらはあくまでも基本的な操作の範疇であり、ある程度 GitHub を使い慣れた者はほとんどの項目に「使いこなしている/十分理解している」と回答する可能性がある。より熟練度が高い場合の評価項目もあり得る。

2.2 アンケート結果からの判定方法

前節で述べたアンケートの結果から習熟度を判定する方法について述べる。

1. 点数化

まずアンケート回答一つ一つの点数化を行う。回答は表 2 の左列の選択肢であるが、これらの回答それぞれに同表右列の点数を付ける。

2. 分類ごと集計

次に表 1 の分類ごとに集計を行う。具体的には「基本 1」5 項目の平均点、「基本 2」4 項目の平均点、「ブランチ 1」5 項目の平均点、「ブランチ 2」6

項目の平均点を計算する。それぞれ最低点は 0.0 点、最高点は 3.0 点となる。

3. 習熟度判定

4 つの平均点から以下の基準で習熟度を判定する。

- 4 つの平均点がすべて 1.0 未満なら「入門者」
- 上記以外で 1.0 未満が 2 分類以上なら「初級者」
- 上記以外で 2.0 未満が 2 分類以上なら「中級者」
- それ以外なら「上級者」

この判定方法は、3.1 で述べる学内ハッカソンやそれ以前の予備調査段階のアンケート結果と、調査対象となった学生の状況を関係教員間で判断して策定したものであり、継続的な評価実験により引き続き洗練していく予定である。

3 評価実験

第 2 節で述べた GitHub 習熟度判定方法の評価実験として、2019 年 2 月に著者ら所属大学で開催された学内ハッカソンの前後と、学部 3 年生向けの必修 PBL 科目の前期開始時と前期終了時にアンケート調査を行い、GitHub 習熟度評価を行った。その結果を述べる。

3.1 学内ハッカソン

2018 年度末の春休み時期、2019 年 2 月に著者ら所属大学で学内ハッカソンが開催された。[8] 著者らもこの開催に協力したが、主要な運営メンバは大学院生を中心とした学生数名であった。プレ PBL としての育成型ハッカソンを銘打ち、学部 1,2 年生のチーム開発未経験者を対象とした初心者部門と、学年等の条件を問わない経験者部門を設定し、2018 年 10 月の約一か月間にチーム単位で参加募集を行った。また同年 12 月には Git/GitHub やアイデア出しなどを対象とした事前勉強会を開催した。参加申し込みは初心者部門 10 チーム、経験者部門 2 チームであったが、経験者部門の 2 チームは日程的な都合が合わず、結果としてハッカソン当日には初心者部門 10 チームのみが参加した。ハッカソン自体は 2 月 16 日～17 日の 2 日間で行われ、発表会や懇親会などの時間を除き、実質

表 1 アンケート項目と分類

項目名	分類
コミットができる	基本 1
コミット間の diff がみれる	基本 2
リモートのリポジトリにプッシュができる	基本 1
リモートのリポジトリから pull ができる	基本 1
リモートのリポジトリのクローンができる	基本 1
git commit -amend ときいて使い方がわかる	基本 2
.gitignore が何のためにあるファイルか知っている	基本 2
コミットをどの程度の粒度にすべきか分かる	基本 2
コミットメッセージにはどんな内容を書くべきか分かる	基本 1
ブランチの作成ができる	ブランチ 1
ブランチの削除ができる	ブランチ 1
あるブランチとあるブランチ 2 の差分がすぐに出せることを知っている	ブランチ 2
あるブランチ A とあるブランチ B を比較して、A のみに加えられた差分をすぐに出せることを知っている	ブランチ 2
merge ができる	ブランチ 1
merge に失敗した後、merge 前に戻るやりかたを知っている	ブランチ 2
rebase ができる	ブランチ 1
rebase に失敗した後、merge 前に戻るやり方を知っている	ブランチ 2
merge 中のコンフリクトを解消できる	ブランチ 2
どんな名前がブランチ名として適しているか分かる	ブランチ 1
どのくらいの頻度 (単位) でマージするのがよいか分かる	ブランチ 2

表 2 アンケート回答の選択肢と点数

選択肢	点数
何のことも分からない	0
コマンドは分かる/使ったことはある	1
分かる/使っている	2
使いこなしている/十分理解している	3

的な開発期間は約 1.5 日程度であった。

GitHub 習熟度判定方法の評価実験として、初心者部門 10 チームの計 48 名 (学部 1 年生 17 名, 同 2 年生 31 名) を対象としてアンケート調査を行ったが、ハッカソン前後共に回答を得られたのは 14 名のみであった。14 名の評価結果を表 3 に示す。表には、ハッカソン前後それぞれのアンケート回答の分類別平均とそれに基づく習熟度判定結果を、ハッカソン前のアンケート回答全体の平均の降順で示している。

表から分かる通り、学生 1,2 はハッカソン前の時点で既に上級者と判定されており、チーム開発初心者部門の参加者であったが、事前に Git/GitHub をそれなりに使っていたことが分かる。こういった学生は、ハッカソン前に行われた Git/GitHub 勉強会で基本的な操作を習得したか、ハッカソンで GitHub を使用することが分かって自習していた可能性がある。また、表中の学生 14 はハッカソン前にはすべての項目の点数が 0.0 であり、Git/GitHub を全く使ったことがなかったが、ハッカソン後の回答では大幅に点数が上がっている。あくまでも自己評価であるため、過大評価の可能性もあるが、本人なりに大きく向上したを感じた結果の回答と考えられる。

表 3 ハッカソン前後の評価結果

	ハッカソン前					ハッカソン後				
	基本 1	基本 2	プラ 1	プラ 2	判定	基本 1	基本 2	プラ 1	プラ 2	判定
学生 1	3.00	2.00	2.80	2.33	上級者	3.00	3.00	3.00	3.00	上級者
学生 2	2.80	1.25	3.00	2.83	上級者	3.00	2.25	3.00	2.83	上級者
学生 3	3.00	1.75	2.60	1.50	中級者	3.00	2.50	2.20	1.83	上級者
学生 4	1.40	1.75	2.00	1.67	中級者	2.00	2.00	2.00	2.00	上級者
学生 5	1.60	1.25	1.00	1.00	中級者	1.40	1.50	1.20	1.00	中級者
学生 6	1.40	0.25	1.00	0.00	初級者	2.00	2.00	1.40	0.33	中級者
学生 7	1.00	0.25	0.40	0.33	初級者	1.40	0.75	1.20	0.00	初級者
学生 8	0.80	0.00	0.80	0.33	入門者	0.80	0.00	0.80	0.33	入門者
学生 9	1.00	0.00	0.80	0.00	初級者	2.00	0.00	1.20	0.00	初級者
学生 10	1.00	0.25	0.60	0.00	初級者	1.40	1.25	1.60	0.67	中級者
学生 11	1.00	0.00	0.40	0.00	初級者	1.80	0.75	1.80	1.17	中級者
学生 12	0.00	0.00	1.00	0.00	初級者	2.80	0.75	2.20	0.67	初級者
学生 13	0.20	0.00	0.00	0.00	入門者	1.60	0.50	1.20	0.33	初級者
学生 14	0.00	0.00	0.00	0.00	入門者	3.00	1.25	2.40	0.50	中級者

3.2 PBL

2018 年度末春休みの学内ハッカソンに続き、2019 年度の学部 3 年生向け必修通年 PBL 科目「プロジェクト学習」[6]においても、評価実験を行っている。通年の PBL であるため、年間通して調査を行う予定だが、プロジェクト開始時(5 月)と前期終了時(7 月)の調査結果について速報する。

PBL 科目「プロジェクト学習」は学部 3 年生全員が履修する必修科目であり、複数教員がチームを組んで提案した 20 数個のテーマに対して、個々の学生の希望などを踏まえてプロジェクト配属が行われ、5 月初旬から夏休み/冬休みを挟んで翌年 1 月末まで続けられる。プロジェクトテーマは活動中にシステム開発を伴うものもそうでないものも様々あるため、システム開発が活動の中心となり、GitHub を使う可能性の高い 3 つのプロジェクト A,B,C に協力を依頼した。各プロジェクトに配属された学生の人数は、プロジェクト A が 15 名、プロジェクト B が 15 名、プロジェクト C が 11 名の計 41 名である。

3 プロジェクトともプロジェクト開始後の 5 月に GitHub 勉強会をそれぞれ開催したが、その勉強会の前の時点で 1 回目のアンケート調査を行い、7 月 19

日に全プロジェクトで行われたプロジェクト中間発表会の後に 2 回目のアンケート調査を行った。1 回目のアンケートは対象者 41 名全員の回答が得られたが、2 回目のアンケートについて本稿執筆時点までに回答が得られたのはプロジェクト A が 15 名、プロジェクト B が 15 名、プロジェクト C が 10 名の計 40 名であった。1 回目、2 回目の両方回答が得られた学生の習熟度判定結果は表 4 の通りである。表には、1 回目アンケートの分類ごと集計結果と判定と、2 回目アンケートの分類ごと集計結果と判定を、所属プロジェクトごとに 1 回目アンケート結果の点数の降順で示している。

この結果から、プロジェクト開始時点でプロジェクト A とプロジェクト C の半数程度、プロジェクト B のほとんどの学生は入門者と判定されていた。どのプロジェクトも前期中には本格的な実装フェーズにはまだ入っていないものと思われるが、前期終了時点の入門者比率はプロジェクト A とプロジェクト B では 2 割から 3 割程度、プロジェクト C では 0%となった。また、上級者と判定されたのはプロジェクト開始時点では各プロジェクトとも 1 名であったが、前期終了時点ではプロジェクト A では 2 名、プロジェクト B で

は3名に増えた。全体的にGit/GitHubに対するスキルアップが進んでいると捉えることができる。一方で、プロジェクト開始時から前期終了時にかけて、判定結果が下がるほどではないものの点数が下がっている学生も見受けられる。分類ごとの点数低下は、アンケート結果そのものの低下であり、習熟度判定方法の問題というわけではない。当該学生が当初は分かっていると思ったものが使っていくうちに実は十分分かっていなかったことに気づいた場合などに起こり得ることであり、主観的評価により判定していることに起因するものと考えられる。

4 考察

4.1 習熟度判定方法の妥当性

アンケートによるGitHub習熟度判定の結果は、教員の立場から見ていた範囲では大よそ妥当に見えるが、教員としても客観的かつ定量的な評価方法を持っているわけではないため、この判定結果が妥当であるかどうかの判断は難しい。そのため、表1の評価項目の妥当性、それらの項目を4つの分類に分ける意味や分け方の妥当性、分類ごと平均からの判定方法妥当性などについて、十分検討する必要がある。

評価項目について、複数のGitHub熟練者で検討して決めたものの、この分け方が妥当であるかどうか引き続き検討を続けるのが望ましいと考えられる。また、より発展的な知識や操作を項目に入れないままで良いのかどうかの検討も必要である。

4つの分類については、当初、基本の項目よりもブランチの項目の方が応用的であり、また、基本とブランチのそれぞれの中で早くに身につく項目とそうでない項目とで分けるということを考えていた。つまり、同じ分類の中の項目は近いタイミングで点数が向上し、別の分類の項目の点数はそうではないということ想定していたが、実際にそうなっているかどうかの評価を行っていない。

更には、分類の当初の意図からすると、少なくとも基本1と基本2、ブランチ1とブランチ2は習熟度評価に対する重み付けが異なるはずのものであるが、現状それを考慮した習熟度判定方法にはなっていない。2.2節で述べた判定方法は、学内ハッカソン前に

実施したGitHub勉強会の際の予備調査や、学内ハッカソン前後のアンケート結果から、大よそこの判定方法で判定できそうであると判断して決めた方法であり、項目の分類わけの主旨も踏まえつつ、習熟度判定方法全体の再検討も必要であろう。

4.2 学生は正直に答えるか

本提案の評価手法はアンケート調査に全面的に依存するものであり、つまり、対象学生がアンケートに正直に答えているかどうか依存している。現状、学内ハッカソンの結果やPBLの成績に直接関わらない形でアンケートを実施しているため、故意によく答えるなどの動機になる要素はあまりないが、単にアンケートに答えるのが面倒だからと適当に回答するという可能性は十分あり得る。また、そもそもの目的はPBLの評価に活用することであるため、そういう運用を始めた時点で学生は自分に都合の良い回答をする可能性は非常に高いと言える。

ただし、3.2節で述べた「プロジェクト学習」には前期末、後期末にそれぞれ学生が自己評価をするフェーズがあるため、直接成績評価に用いるのではなく、自己評価の一部として本提案の評価手法のようなアンケートを実施するという利用方法はあり得る。

4.3 客観的評価へ

本提案の方法はあくまでも主観的な評価であるが、理想的には同じものを客観的に評価したい。第1節で述べた通り、GitHubには相応に利用履歴が蓄積され、後で参照することが可能であることから、利用履歴データによる客観的評価を目指したい。ただし、本提案手法に挙げたような項目すべてに直接対応するデータを利用履歴のみから取得するのは困難であることが予想される。そのため、本提案手法による主観的評価の結果と、GitHub利用履歴から取得できるデータによる評価結果の関連性などを分析して、全体として評価の客観性を高められるのが望ましいと考えている。

表 4 PBL 開始時/前期終了時の評価結果

	所属	プロジェクト開始時					前期終了時				
		基本 1	基本 2	ブラ 1	ブラ 2	判定	基本 1	基本 2	ブラ 1	ブラ 2	判定
学生 20	A	3.00	2.50	2.60	1.83	上級者	3.00	1.75	2.60	2.00	上級者
学生 21	A	1.80	1.25	1.20	0.67	中級者	2.00	1.50	1.60	1.33	中級者
学生 22	A	1.40	1.25	1.00	1.00	中級者	2.80	2.75	3.00	2.83	上級者
学生 23	A	1.80	0.50	1.40	0.50	初級者	2.00	1.25	1.80	1.17	中級者
学生 24	A	1.80	1.00	0.80	0.50	初級者	1.80	1.50	1.80	1.33	中級者
学生 25	A	2.00	0.50	1.20	0.33	初級者	1.00	0.25	1.60	0.67	初級者
学生 26	A	1.60	0.50	1.00	0.00	初級者	1.60	0.25	0.20	0.00	初級者
学生 27	A	0.80	0.25	0.40	0.00	入門者	2.00	1.25	1.60	0.67	中級者
学生 28	A	0.20	0.00	0.40	0.00	入門者	1.80	0.50	1.20	0.50	初級者
学生 29	A	0.20	0.00	0.20	0.00	入門者	0.80	0.25	0.80	0.33	入門者
学生 30	A	0.20	0.00	0.00	0.00	入門者	1.20	0.00	0.20	0.00	初級者
学生 31	A	0.00	0.00	0.00	0.00	入門者	0.60	0.00	0.80	0.33	入門者
学生 32	A	0.00	0.00	0.00	0.00	入門者	1.20	0.00	0.60	0.17	初級者
学生 33	A	0.00	0.00	0.00	0.00	入門者	2.00	0.25	0.80	0.33	初級者
学生 34	A	0.00	0.00	0.00	0.00	入門者	0.20	0.00	0.00	0.00	入門者
学生 35	B	3.00	3.00	2.40	2.67	上級者	3.00	3.00	3.00	3.00	上級者
学生 36	B	2.00	1.50	1.60	0.50	中級者	3.00	2.50	2.60	2.33	上級者
学生 37	B	2.00	1.00	1.60	1.00	中級者	2.40	1.25	1.60	1.17	中級者
学生 38	B	2.00	0.25	1.60	0.33	初級者	3.00	2.25	2.60	2.17	上級者
学生 39	B	0.80	0.50	0.60	0.00	入門者	2.00	1.25	1.80	1.33	中級者
学生 40	B	0.60	0.25	0.20	0.00	入門者	2.00	0.50	0.80	0.33	初級者
学生 41	B	0.00	0.00	0.00	0.00	入門者	1.60	0.75	1.20	0.83	初級者
学生 42	B	0.00	0.00	0.00	0.00	入門者	0.60	0.00	0.80	0.00	入門者
学生 43	B	0.00	0.00	0.00	0.00	入門者	0.00	0.00	0.20	0.00	入門者
学生 44	B	0.00	0.00	0.00	0.00	入門者	2.00	0.00	1.20	0.00	初級者
学生 45	B	0.00	0.00	0.00	0.00	入門者	0.80	0.50	0.20	0.00	入門者
学生 46	B	0.00	0.00	0.00	0.00	入門者	1.20	0.75	0.80	0.50	初級者
学生 47	B	0.00	0.00	0.00	0.00	入門者	1.00	0.00	0.40	0.00	初級者
学生 48	B	0.00	0.00	0.00	0.00	入門者	1.20	0.25	0.60	0.00	初級者
学生 49	B	0.00	0.00	0.00	0.00	入門者	1.00	0.75	1.00	0.83	初級者
学生 50	C	2.80	2.50	2.80	1.67	上級者	2.60	2.25	2.00	2.00	上級者
学生 52	C	1.80	0.50	1.00	0.83	初級者	2.00	1.50	2.00	1.67	中級者
学生 53	C	2.00	0.25	0.80	0.33	初級者	1.00	0.50	0.60	0.33	初級者
学生 54	C	1.00	0.25	0.60	0.17	初級者	1.80	1.25	2.00	1.17	中級者
学生 55	C	0.80	0.00	0.60	0.17	入門者	1.00	0.25	1.00	0.17	初級者
学生 56	C	0.80	0.25	0.40	0.00	入門者	2.00	0.50	1.40	0.67	初級者
学生 57	C	0.80	0.00	0.20	0.00	入門者	1.80	0.50	1.60	0.83	初級者
学生 58	C	0.40	0.00	0.00	0.00	入門者	1.20	0.25	0.80	0.17	初級者
学生 59	C	0.00	0.00	0.00	0.00	入門者	1.80	0.50	1.80	1.17	中級者
学生 60	C	0.00	0.00	0.00	0.00	入門者	1.00	0.25	0.80	0.17	初級者

5 おわりに

システム開発 PBL に参加する学生の GitHub 習熟度の評価基準および評価手法を提案し、実際に学内ハッカソンや PBL において継続的に評価実験を行っている。現状、アンケートによる主観的評価に依るものであり、十分に洗練されたものではないものの、学生の GitHub 習熟度を比較的低コストに評価できる状況になってきている。引き続き評価実験を行いながら、評価手法の洗練させるのと並行し、GitHub 利用履歴データ等を利用した客観的評価を用いた評価手法も検討していきたい。

参考文献

- [1] enPiT 事務局: 高度 IT 人材を育成する産学協働の実践教育ネットワーク, 2017.
- [2] 福田晃, 鶴林尚靖, 荒木啓二郎, 峯恒憲, 日下部茂, 金子邦彦, 亀井靖高, 廣重法道: 情報工学系大学教員のための *PBL* 実践ガイド, 九州大学大学院システム情報科学府 情報知能工学専攻社会情報システム工学コース, 2012.
- [3] 日戸直紘, 伊藤恵, 大場みち子: 能力成熟度モデル統合に基づいた PBL における定量的学習評価手法の提案, 日本ソフトウェア科学会第 34 回大会論文集, 日本ソフトウェア科学会, 2017.
- [4] @kyanro@github: git 理解度チェックリスト, <https://qiita.com/kyanro@github/items/3f0d4303882e668e1e85> (2019 年 8 月 1 日アクセス).
- [5] 松原克弥, 伊藤恵, 木塚あゆみ: コード更新履歴の可視化による分散 PBL でのチーム活動評価の試み, 第 4 回実践的 IT 教育シンポジウム (*rePiT2018*) 論文集, 日本ソフトウェア科学会, 2018, pp. 49–55.
- [6] 公立はこだて未来大学: プロジェクト学習, http://www.fun.ac.jp/edu_career/project_learning/ (2019 年 8 月 1 日アクセス).
- [7] 大塚弘記: *GitHub* 実践入門, 技術評論社, 2014.
- [8] P2HACKS 実行委員会: P2HACKS – プレ PBL としての育成型ハッカソン, <https://p2hacks.c.fun.ac.jp/2019/> (2019 年 8 月 1 日アクセス).